JOURNAL OF RESEARCH of the National Bureau of Standards—A. Physics and Chemistry
Vol. 75A, No. 5, September–October 1971

Discrimination Between Equations of State*

J. Ross Macdonald** and Donald R. Powell**

(June 16, 1971)

Eight isothermal equations of state are analyzed to yield quantitative measures of the degrees to which equation pairs can be discriminated for real data, data of limited span and precision. Calculated curves allow one to assess the span and precision necessary in P-V data to allow unambiguous discrimination of various pairs. Some discussion is presented of bias and systematic error which may arise in least squares fitting. Using exact synthetic data, we also illustrate for seven equation pairs the very large relative systematic errors in parameter and standard deviation estimates which arise from such fitting of data of limited span with an incorrect but "close" equation model. General conclusions following from these results are discussed. Although the present work is principally concerned with discrimination between equations of state, its results are pertinent to the more general problem of choosing a "best" analytical model (linear or nonlinear) to represent experimental results.

Key words: Curve fitting; data analysis; equations of state; model discrimination; surface fitting.

1. Introduction

Virtually all physical science is concerned at some stage with comparing experimental data with theoretical predictions. Although no theories are ever fully verifiable, one nearly always wants to find that theoretical model, from the limited set of possible models under consideration, which best represents the data, which allows the underlying phenomena to be better understood, and, if possible, which allows prediction outside the range of the original measurements. In the relatively early stages of investigation of a given domain, one usually does not know which of several theoretical or empirical models is likely to be most appropriate. This state of affairs is particularly likely to occur when the physical situation being studied is too complex to allow a tractable theoretical idealization, which is still sufficiently close to the experimental situation, to be accurate. Many-body interaction problems, such as that of determining the exact equation of state of a solid or liquid, fall in this category.

The problem of model discrimination is made difficult by the presence of random and systematic errors in the data. In the present paper, it is assumed that systematic error in the data is absent or at least negligible compared with other error. Systematic error can still be generated, of course, by the choice of an inappropriate model [1],[1] and a question of considerable importance is: Under what conditions is it possible to discriminate adequately between several more-or-less-appropriate models, or equations? In the present paper, we shall be concerned with typical synthetic equation-of-state data generated without significant error of any kind, reserving a detailed discussion of the effect of random errors to a later paper. It will be shown that by using such exact "data" we can investigate what sort of discrimination is possible between various equations of state in practical cases where measurements are of limited precision.

In real life, experimental data have only limited accuracy and precision and always extend only over a limited range of the variables involved. This state of affairs suggests intuitively that one will be unable to discriminate adequately between two or more analytical models which are sufficiently close together in their predictions for the range considered. We are here concerned with ways of making this intuition quantitative at least for the specific equations considered here. Since better discrimination may sometimes appear possible than is actually the case, just because of the presence of more or less random errors which happen to fall in a particular way, it is important to consider exact data before data with random errors.

Although all that is often required of an equation of state, or more generally, a mathematical model of experimental results, is that it serve adequately as an interpolation and smoothing device for the data, the problem of model discrimination is usually still present even in this case. Unless the first model fitted passes all tests of adequacy, more than one model must be examined and a choice of available models made. The present paper discusses some general methods of model discrimination with specific illustrations taken from the equation of state field. Here we are concerned additionally with the task of estimating physically significant parameters of the material which led to the data in question.

Two somewhat different situations frequently arise in the equation of state area. Often one starts with no,

*An invited paper.
**Present address: Texas Instruments Incorporated, P.O. Box 5936, Dallas, Texas 75222.
[1] Figures in brackets indicate the literature references at the end of this paper.

or only crude, knowledge of the underlying parameters of the material under investigation. These parameters are then determined by fitting various equations of state to P-V data, usually by least squares techniques [1]. The most appropriate, or "best fit" equation will usually be that which leads to minimum estimated standard deviations of the fitted data points and of the parameters. The values of the parameters obtained from this fit are taken to be the best available estimates of the unknown material parameters. In general, however, such values will *not* usually be good estimates unless the choice of model is appropriate for the data and leads to randomly distributed, essentially stochastically independent residuals, and the fitting procedure itself leads to negligibly biased parameter estimates.

Sometimes one is able to obtain estimates of the parameters by other means than from fitting of direct P-V measurements. Now becoming popular for this purpose is the method of ultrasonic velocity measurements under pressure [2, 3], pioneered by Lazarus [4]. Once having parameter estimates available for a certain material, one can, given the appropriate equation of state, calculate volume values for a range of pressure. Of course, with a limited number of parameters available, as is always the case, calculated volumes can generally only be expected to remain reasonably accurate over a limited range of pressure. The nub of the problem here is usually in knowing what equation of state to use (and how far to trust it). Sometimes the determination of the best equation may be made by comparing ultrasonically derived parameters with those obtained from a least squares fit of direct P-V data for the same material.

In either approach, one eventually obtains a set of parameters believed to be appropriate for the material under investigation. Although in actual practice these parameters will always be uncertain to some degree, it is nevertheless useful to ask, as a limiting case, how well one can distinguish between various equations of state when the parameters are actually exact (or are so considered) but when available P-V data are of limited precision. Some answers to this question are discussed later for eight different equations of state of some current interest.

One of the important purposes of the present work is to point out that uniqueness is a limit seldom achieved in practice. Frequently an experimenter chooses a model to represent data of given range with the implication or statement that the chosen model is "best" or "most applicable" without realizing or investigating sufficiently to find that other different models are equally applicable for the given data.

Although the present analysis is concerned with discrimination between eight specific equations of state and thus involves quantitative results only for these equations, we expect that the results will also apply at least qualitatively to other not-too-different equations. More importantly, perhaps, the present discrimination methods and general approach can and should be applied to any experimental situation where it is important to establish one or more adequate mathematical representations of the data or, better, of the underlying process which led to the predictable part of the data.

2. Equations of State Considered

The material parameters with which we shall be concerned, all for isothermal conditions, are the specific volume, V_0, at a given reference pressure P_0; the bulk modulus at $P=P_0$, $K_0 \equiv -V_0(\partial P/\partial V)\,|_{P=P_0}$ and various pressure derivatives of the bulk modulus, K, also evaluated at $P=P_0$. For simplicity, let $p \equiv P-P_0$; then $V=V_0$ at $p=0$. Now $K_0' \equiv \eta \equiv (\partial K/\partial P)\,|_{p=0}$, and $K_0'' \equiv (\partial^2 K/\partial P^2)\,|_{p=0}$. The symbol η has been introduced to simplify subsequent equations; it is dimensionless. It is also useful to introduce the further dimensionless quantity $\psi \equiv K_0 K_0''$. Finally, define the dimensionless pressure variable $z \equiv p/K_0$ and the dimensionless density variable $x \equiv \rho/\rho_0 \equiv V_0/V$.

Barsch and Chang [3] have recently given values for the parameters of CsI at 25 °C, plus temperature derivatives of these parameters. The quantity V_0 may be calculated from x-ray measurements of the lattice constant. Other parameters such as K_0, K_0', and K_0'' were obtained from ultrasonic measurements. Using the Barsch and Chang results, we have calculated the values of V_0, K_0, η and ψ which then apply to CsI at 150 °C, an arbitrary choice of temperature. These values, as used in our computer studies, have 14 figure accuracy and may be considered the accurate values of some hypothetical material close to CsI at 150 °C. Of course as applied to CsI itself, only a few places in each parameter value are significant. To five figures, the parameter values are: $V_0 \cong 1.0184$, $K_0 \cong 1.0503 \times 10^2$ kbar, $\eta \cong 6.0382$, and $\psi \cong -6.9897$. Here we have taken $P_0=0$ and V_0 at 25 °C as unity. Thus, all volumes used here are reduced specific volumes and are dimensionless. The original Barsch and Chang 25 °C values are $V_0=1$, $K_0=118.9\pm0.6$ kbar, $\eta=5.86\pm0.11$ and $K_0''=-0.052\pm0.002$ kbar^{-1}. These results lead to $\Psi \cong -6.2$ at 25 °C.

We shall be interested here in comparisons of, or discrimination between, eight different equations frequently employed in equation of state studies [1, 3]. We have adopted the approach of Barsch and Chang of designating the ordinary Murnaghan equation as the first-order Murnaghan equation (ME_1), and the equation previously [1] termed the second-order equation (SOE) as the second-order Murnaghan equation (ME_2). There are several forms of this latter equation, depending upon the values of η^2 and Ψ; here only one of these forms is pertinent. All eight of the equations are given in the form $z=f(x)$ in table 1, which also lists acronyms for each equation. Some, but not all of them, may be expressed in inverse form, with x as an explicit function of z. Note that three of the equations are "first-order" in the sense of Barsch and Chang [3]. They involve $n=3$ parameters: V_0, K_0, and η. The other five "second-order" equations involve Ψ in addition. Finally, table 1 includes values of $K_\infty' \equiv (\partial K/\partial P)_{p\to\infty}$.

TABLE 1. *Equations of state of interest written in the dimensionless form* z = f(x)

Equation	Acronym	Form	n	K'_∞
		$z \equiv p/K_0 = f(x)$ $\quad f(x)$ $\quad x \equiv \rho/\rho_0 \equiv V_0/V$		
Usual Tait	UTE	$(\eta+1)^{-1}[\exp\{(\eta+1)(1-x^{-1})\}-1]$	3	$-\infty$
First-order Murnaghan	ME_1	$\eta^{-1}[x^\eta-1]$	3	η
Second-order Murnaghan	ME_2 $(\eta^2 \geq 2\psi)$	$2(x^{(\eta^2-2\psi)^{1/2}}-1)/[(\eta^2-2\psi)^{1/2}(x^{(\eta^2-2\psi)^{1/2}}+1) - \eta(x^{(\eta^2-2\psi)^{1/2}}-1)]$	4	$-\infty$
Keane	KE $(-\eta^2 < \psi < 0)$	$[\eta^3/(\eta^2+\psi)^2][x^{(\eta^2+\psi)/\eta}-1]-[\psi/(\eta^2+\psi)]\ln x$	4	$(\eta^2+\psi)/\eta$
First-order Birch	BE_1	$(3/2)[x^{7/3}-x^{5/3}][1+(3/4)(\eta-4)(x^{2/3}-1)]$	3	3
Second-order Birch	BE_2	$(3/2)[x^{7/3}-x^{5/3}][1+(3/4)(\eta-4)(x^{2/3}-1) + (1/24)\{143+9\eta(\eta-7)+9\psi\}(x^{2/3}-1)^2]$	4	11/3
Third-degree Slater	3SE	$(1-x^{-1})+(1/2)(\eta+1)(1-x^{-1})^2+(1/6)(\eta^2 + 3\eta+2+\psi)(1-x^{-1})^3$	4	-3
Third-degree Davis-Gordon	3DGE	$(x-1)+(1/2)(\eta-1)(x-1)^2 + (1/6)(\eta^2-3\eta+2+\psi)(x-1)^3$	4	3

The Keane equation only applies when $-\eta^2 < \Psi < 0$, conditions satisfied for the present parameter values. Although all of the equations must become poor models for sufficiently high z, failure is particularly evident for the UTE and ME_2. The volume predicted by the UTE goes through zero at the finite z value of $(\eta+1)^{-1}[\exp(\eta+1)-1]$. For the present form of the ME_2, $K' \equiv (\partial K/\partial P)=0$ at $z=-\eta/\Psi$ and $V=0$ at $z=2/[(\eta^2-2\Psi)^{1/2}-\eta]$. The 3SE also suffers from the disadvantage that it predicts zero volume at finite pressure. All the other equations require infinite z to produce zero volume.

The equations of table 1 are discussed in greater detail elsewhere [1, 3]. Although most of them have some macroscopic or phenomenological theoretical justification, here they may simply be regarded as empirical equations likely to be of some value in the P-V area.

3. Model Differences and ΔV Discrimination

In order to examine differences in the predictions of the various models, we have, for a given set of p or z values, calculated corresponding dimensionless V values, using in each equation the same 14-figure parameter values already mentioned. The V values were calculated using 14 figures, by iteration when necessary, with a resulting 13-figure or better accuracy. Finally, differences between V values of each possible pair of equations were calculated for each z value. The differences obtained for $p=11$ kbar, or $z \cong 0.1047$, are listed in table 2, all multiplied by 10^4 for convenience. The ΔV's shown are formed by taking the V of one of the equations listed in the left column and subtracting from it the V calculated using one of the equations in the top row. Since the ME_1-3SE ΔV value is largest of all, the ME_1 yields the largest and the 3SE the smallest V value for this value of z. Similarly, we see that the BE_2 and KE volume predictions are closest together here.

In addition, in figures 1 to 5 we have plotted ΔV versus z for a variety of equation pairs. The boxed equation name is the equation from whose V values those of the equations named on the curves are subtracted. These five figures contain ΔV curves for most, but not quite all, of the possible pairs of equations. Curves have not been duplicated. Thus $(V_{BE_2}-V_{BE_1})$ appears in figure 3 for BE_2 but not its negative in figure 5 for BE_1. Negative values are indicated by using dashed lines.

TABLE 2. *Scaled volume differences,* $10^4\Delta V$, *for equation pairs at* $z \cong 0.1047$

Equations	3SE	BE_1	ME_2	UTE	3DGE	BE_2	KE	ME_1
3SE	0							
BE_1	2.1	0						
ME_2	2.9	0.77	0					
UTE	3.7	1.6	0.81	0				
3DGE	3.9	1.8	1.1	0.27	0			
BE_2	4.2	2.1	1.3	0.49	0.22	0		
KE	4.3	2.2	1.4	0.62	0.35	0.14	0	
ME_1	8.9	6.8	6.0	5.2	4.9	4.7	4.6	0

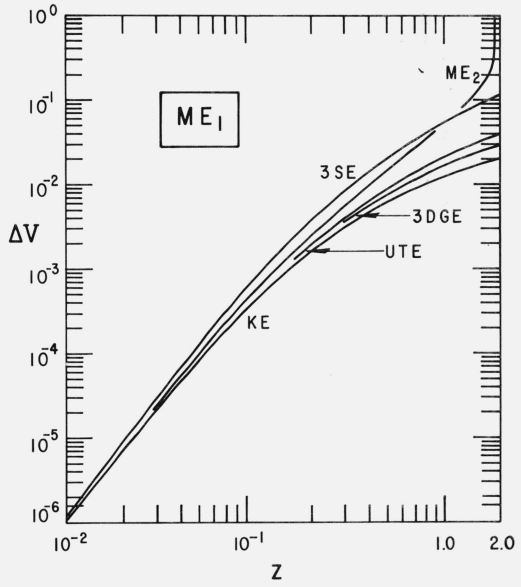

FIGURE 1. *Volume differences, ΔV, versus normalized pressure, z, for the ME_1 and other equations.*

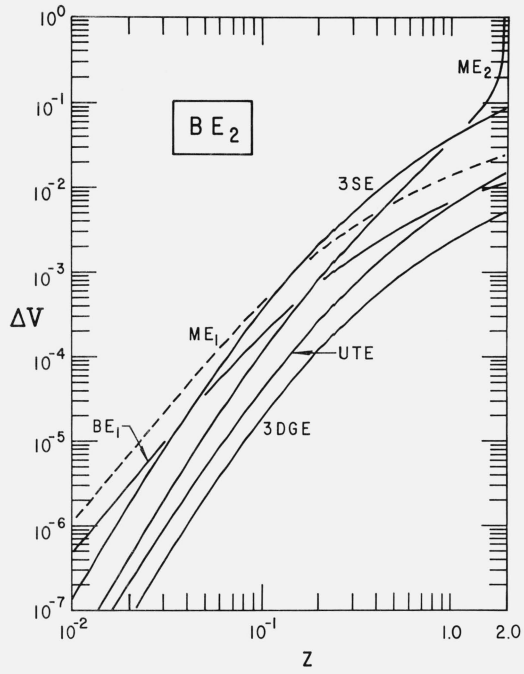

FIGURE 3. *Volume differences, ΔV, versus normalized pressure, z, for the BE_2 and other equations.*

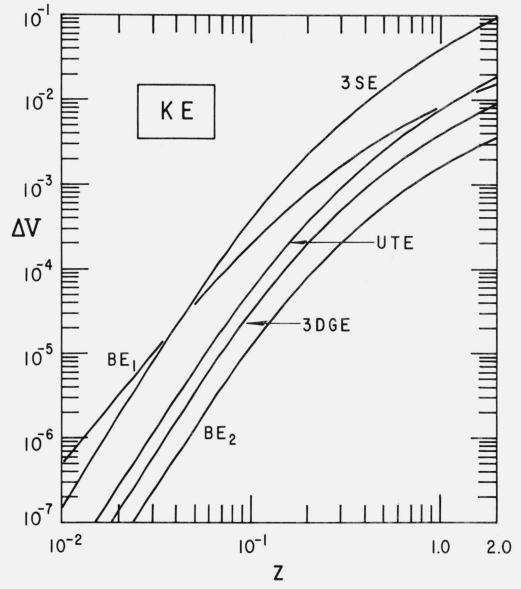

FIGURE 2. *Volume differences, ΔV, versus normalized pressure, z, for the KE and other equations.*

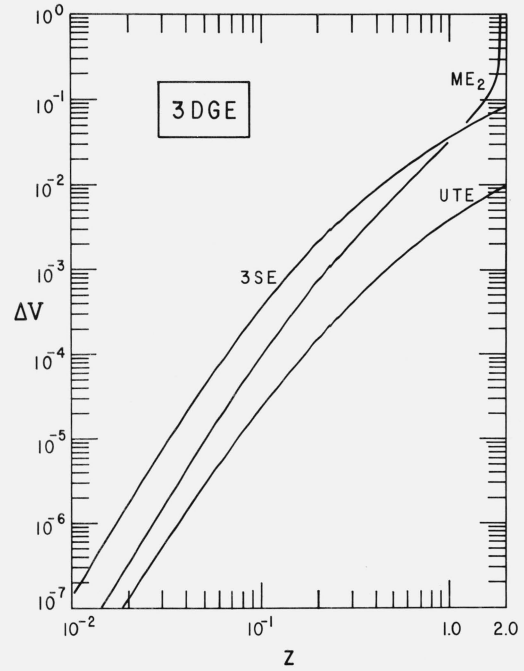

FIGURE 4. *Volume differences, ΔV, versus normalized pressure, z, for the 3DGE and other equations.*

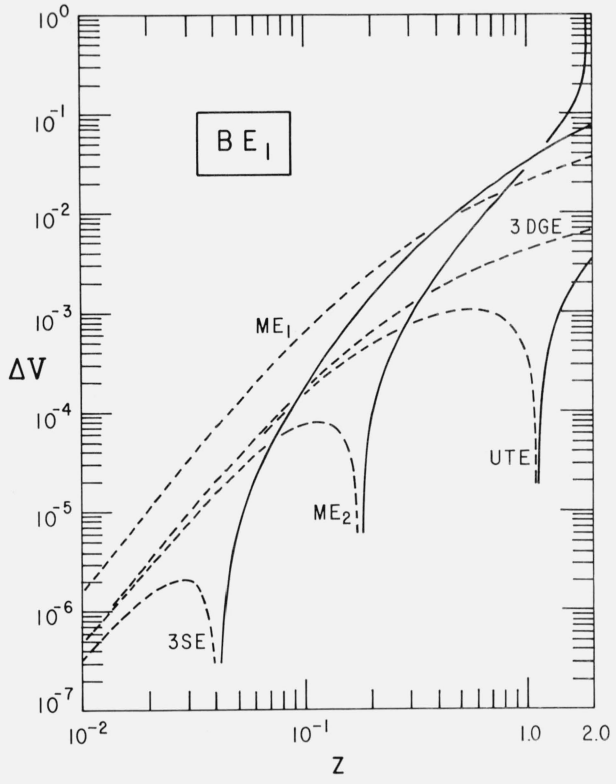

FIGURE 5. *Volume differences, ΔV, versus normalized pressure, z, for the BE_1 and other equations.*

Although few actual experiments resulting in P-V values of appreciable accuracy extend past $z \sim 0.5$, the present exact, synthetic data curves are calculated up to $p = 210$ kbar, where $z \cong 2.000$. At this z value, $V/V_0 \equiv x^{-1}$ is of the order of 0.5 for these equations, being ~ 0.64 for the 3DGE, for example. For $z_{max} \cong 2.000$, $N = 58$ p or z values, distributed roughly logarithmically, were used. For present purposes, larger z values were unnecessary.

Clearly, ΔV curves for all pairs not involving the UTE, 3SE, or ME_2 will eventually reach a maximum, with $\Delta V_{max} < 1$, as z increases, then decrease toward zero since both V's become arbitrarily small as $z \to \infty$. As the figures show, the situation is different for the ME_2 even within the present range. Since the parameter values used here lead to $V < 0$ for $z \gtrsim 1.85$, ΔV values which involve ME_2 volumes become arbitrarily large in magnitude as z increases beyond this point. Clearly, the ME_2 cannot be a useful model all the way to the point where it predicts zero or negative volumes. Nevertheless, it may be useful for a range ending sufficiently far below this point.

Of what value are the results shown in figures 1 to 5? They are of considerable value because they show how well the various equations of state considered here may be discriminated under the best possible conditions. Suppose, for example, that we wish to discriminate between the KE and other equations and are able to measure volume only up to $z = 0.1$. Further, suppose that errors in p are negligible compared to those in V. Figure 2 then shows that to distinguish the 3SE from the KE in the range $0 \lesssim z \lesssim 0.1$, experimentally determined V values must be known to about one part in 10^4, or to four decimal places, near $z \sim 0.1$. Even less uncertainty would be required for a smaller range. The BE_2 and KE cannot be reliably distinguished without a precision of about three parts in 10^6 near $z = 0.1$ and higher precision for smaller z. Clearly, if the above precision has not been achieved, there would be no point in attempting to discriminate between the equation pairs discussed for the data in question. Barsch and Chang [3] have discriminated between the BE_2 and KE for a situation where $\Delta V/V_0 \simeq 3 \times 10^{-3}$ or more and have concluded that the BE_2 was much better for their purposes than the KE. The present figure 2 results indicate that such discrimination is actually not significant with such precision in ΔV, for the present set of parameter values, over a pressure range from zero up to at least 200 kbar.

There are two reasons why we consider that the present curves represent the best possible discrimination. First, there are always some random errors in the determination of pressure values. To first order, we may take the expected or "controlled" pressure values as exact and consider that the actual pressure errors are incorporated as additional random errors in the volume values. It is then this total volume error which must be used in determining whether the curves allow equation discrimination within a certain range of z. When parameter values are available, as from ultrasonic measurements, they may be used in several equations of state to calculate exact volumes over a given z range. These volumes may then be directly compared with a set obtained by direct measurement. Clearly, if the total errors in the latter set are not appreciably smaller (over most or all of the z range) than the ΔV's obtained with various equation pairs, no valid discrimination is possible. Even so, one of the several equations among which discrimination is impossible for the given z range may be far superior to the others for extrapolation beyond this range. Although all eight equations of figures 1 to 5 are indistinguishable for ΔV data of no better than 10^{-3} precision in the range $0 \leq z \leq 0.1$, clearly there are important differences between the predictions of the various equations for this same precision level at say $z = 1.5$.

When an independently measured set of parameter values is unavailable, parameter value estimates must be obtained by fitting a model to the available data by some such procedure as least squares. Each different model fitted will then yield a different set of estimated parameter values. If ΔV values are obtained for a pair of models, using in each model the specific parameter values determined from a least squares fit of the data for the given model and range (case A), then the adjustment of the parameter values associated with the least squares procedure will generally lead to an appreciably different set of ΔV values than would have been obtained had the *same* parameter value set been used in each equation (case B). If the fits of the two equations for case A are sufficiently good, the corresponding ΔV values may nearly all be much smaller than those

obtained in case B with any *single* reasonable parameter value set. But only one, at most, of the two sets of parameter values can be correct. Thus, one must proceed with extreme caution, and the small degree of discrimination possible from the case A fits and ΔV's is usually misleading. Further, any use of case A results outside of their fitting ranges is extremely dangerous.

The most meaningful discrimination will be obtained from calculating ΔV's by the case B procedure, using the same most reasonable choice of parameter values in both equations. If the two equations under consideration seem to fit about equally well and no other parameter value information is available, reasonable values to use in the case B discrimination are the averages of the two sets of values found from the least squares fittings. Because of the wide use of least squares procedures, these matters will be further discussed in the next section.

The present case B results are closely related to some obtained by Barsch and Chang [3]. These authors compared, however, p-value predictions obtained from a certain lattice equation of state tailored for CsI with p values obtained from several other phenomenological equations of state. They found, for example, that using the same parameter values the BE_2 approximated the lattice equation an order of magnitude better (in Δp) than did the KE. Although Barsch and Chang present V/V_0 versus p curves for several of the equations of state considered herein, they do not give ΔV curves and are not primarily concerned with establishing what accuracy in V is needed, for a given p or z range to allow equation discrimination.

Even though Barsch and Chang assert the superiority of the BE_2 over the other phenomenological equations they considered, as already mentioned the BE_2 curve of the present figure 2 suggests that exceptionally accurate data or a very wide range will usually be required to allow meaningful discrimination to be made between the BE_2 and the KE. Although Barsch and Chang's calculated $|\Delta p|$ values for the BE_2 and lattice equation were an order of magnitude smaller than those for the KE and lattice equation, the latter values themselves were still considerably smaller for the range $0 \leq p \leq 200$ kbar than either the errors in $|\Delta p|$ calculated from the BE_2 with experimental uncertainties in the parameters or those expected experimentally [3]. Thus, the actual discrimination between the BE_2 and KE appears nugatory for this range. It seems doubtful that sufficiently accurate wide-range data yet exist to make adequate BE_2-KE discrimination possible unless the situation is very different for appreciably different parameter values than those used here and those used by Barsch and Chang, an unlikely possibility.

The curves of figures 1 to 5 are somewhat more general than they appear at first sight. First, since the normalized pressure variable z is used, the results are independent of the value of K_0. Second, since the V_0 value used is quite close to unity, little specialization is introduced by the specific V_0 value used. When V_0 differs appreciably from unity, the present curves may still be used with the ΔV values reinterpreted as $\Delta V/V_0$ values. For the UTE, ME_1, and BE_1, only the additional parameter η enters. This quantity is usually found to be within the range $3 \lesssim \eta \lesssim 8$; thus, the present value, near 6, is fairly typical. Further, changes in η may be expected to change ΔV itself less than the V's entering ΔV. On the other hand, the Ψ value used, near -7, is quite special since little is known thus far about the likely range of Ψ for a variety of materials, and it probably can be positive as well as negative [1, 5]. Nevertheless, we suggest that the present curves may be used, at a zero to first order level, for an initial estimate of discrimination possibilities between various equations for other materials besides CsI at 150 °C. Of course, the next order of assessment would employ an estimated parameter set (V_0, K_0, K_0', and K_0'' values) for the material in question. This set could then be used, as herein, to generate ΔV curves for comparison with the estimated total errors of the experiment, all incorporated into the V values.

As examples of such zero-order assessment, we may consider the data of Monfort and Swenson [6], Kell and Whalley [7], and Vedam and Holton [8]. Monfort and Swenson studied potassium metal up to $z \sim 0.4$. Their volume data were given to four places, and they found a scatter of about 5 units in the last place. Although they primarily considered the BE_1, the ME_1 was also introduced. The ME_1 curve of figure 5 shows a maximum $|\Delta V|$ for these equations of about 7×10^{-3}. When the Monfort-Swenson data is normalized to a V_0 near unity, allowing comparison of V errors with present ΔV's, one may estimate that the data are accurate to perhaps 3×10^{-3} in normalized volume. Comparison suggests that one might then just be able to distinguish between the BE_1 and ME_1 for this range and accuracy. One of the present authors [1] has considered discrimination between the 3DGE and 3SE for the 0 °C water data of Kell and Whalley ($z_{max} \sim 0.05$) and between the 3DGE and ME_2 for the 50 °C water data of Vedam and Holton ($z_{max} \sim 0.44$). Similar zero-order comparison of probable errors in V with the present ΔV curves suggests that the 3DGE–3SE discrimination was near the borderline of possibility and was probably not very meaningful, while the 3DGE–ME_2 discrimination was somewhat more possible and certain.

Finally, to the degree that the present ΔV curves are reasonably general, it is worth mentioning that the sign changes for the $V_{BE_1} - V_{UTE}$, $V_{BE_1} - V_{ME_2}$, and $V_{BE_1} - V_{3SE}$ curves shown in figure 5 indicate that the BE_1 remains a closer approximation to the other three equations over a wider range than if such changes of sign were absent. This result is perhaps one reason why the BE_1 has been found to be of relatively general applicability in the past.

4. Least-Squares Comparisons

Least squares procedures are frequently applied to noisy data for which the true underlying model is unknown and possibly nonlinear in some of the parameters. Here we shall investigate the results of least squares fitting of exact data, especially with incorrect models. Such analyses, when the correct model and parameter values which generated the data are known,

can yield valuable information about the systematic errors arising from the use of the wrong model. Further, the use of exact data allows the usually mixed effects of random and systematic errors of this type to be entirely separated. Since figures 1 to 3 and 5 show that the 3DGE is, in some sense, close in its predictions to several of the other equations, it has been chosen here as the "correct" model for illustrative purposes. The exact data employed was thus generated by using in the 3DGE the 150 °C CsI parameter values already discussed.

Table 3 shows the results of applying the least squares method in a few different situations of interest. Here and hereafter "linear" and "nonlinear" generally refer to linearity, or its absence, of the parameters entering the model. Thus, by a "linear" equation we will mean one linear in its parameters. The "linear" situation cited is actually rendered nonlinear in the parameters by the weighting used [1]. Even though the model is originally linear in the parameters, weighting of the independent variable will lead to nonlinear parameter dependence except in the special simple case (not considered here) of a linear relation between independent and dependent variables. In a succeeding paper, we hope to investigate in some detail the importance of and degree of bias frequently arising in the A case of table 3 when random error is present. Here we continue to restrict attention to the exact data situation.

The 3DGE is written in table 1 in a form involving the parameters nonlinearly. This form was required by the constraint of using K_0, η, and Ψ as the basic parameters in each equation listed in the table. On the other hand, the 3DGE may also be written in the linear form

$$p=\sum_{i=0}^{3} A_i(x-1)^i \qquad (1)$$

where $A_0=0$ when the V_0 entering $x\equiv V_0/V$ is taken fixed and has its correct value (the procedure we shall use when A_0 is a free parameter); $A_1\equiv K_0$; $A_2\equiv(\eta-1)K_0/2$; and $A_3\equiv(1/6)(\eta^2-3\eta+2+\Psi)K_0$. Clearly, direct linear least squares determination of the A_i parameter estimates will allow corresponding V_0, K_0, η, and Ψ estimates to be calculated for comparison with their true values. Further, comparison of corresponding nonlinear and linear least square parameter estimates will allow bias arising from nonlinear least squares to be indentified and quantified.

The following definitions are useful in comparing least-squares parameter values with exact values. Let θ be a specific parameter of the model; then denote the true value of θ (here known) by θ_0 and the least-squares estimate as $\hat{\theta}$. The relative error of the estimate is then $\delta\equiv(\hat{\theta}-\theta_0)/\theta_0$. When no random errors are present, δ_i will measure the systematic error in the ith parameter value. It is also of interest to compare the parameter error $(\hat{\theta}-\theta_0)$ with the standard deviation $(s_d)_\theta$ obtained for a given least squares estimate of θ. To do so, we define $\Delta\equiv|(\hat{\theta}-\theta_0)/(s_d)_\theta|=|\theta_0\delta/(s_d)_\theta|$. This measure will indicate possible systematic error in $(s_d)_\theta$.

We have been discussing least squares results in the above as though they were exact solutions of the least squares equations. It is not widely appreciated that all the usual least squares computer routines may yield very inaccurate parameter values because of round-off errors [9]. For example, if Gaussian elimination with pivoting is used to solve the least squares equations, the number of accurate decimal digits in a $\hat{\theta}$, A, is $A\sim(C-n+1\pm1)$, where C is the number of (equivalent) decimal digits carried in the computer calculation and n is the number of free parameters. Clearly, if $n\gtrsim C$, results of little value are likely to be obtained. Expression for A of this type were originally derived for linear least squares fitting of polynomials, but they seem to apply at least approximately to nonlinear equations as well. Recently, Wampler [10] has made a more detailed study of the matter for polynomials and discussed more complex routines which can yield very substantially higher solution accuracy.

The effects of roundoff are illustrated by the results of table 4. Elimination with pivoting was used to carry out least squares fitting of the 14-figure 3DGE data using the 3DGE equation in both its linear and nonlinear forms. Since $c=14$ *and* $n=4$, $A\sim11\pm1$. In Table 4, the δ_i are calculated using $\theta_0=V_0$, $\theta_1=K_0$, $\theta_2=\eta$, and $\theta_3=\Psi$. The quantity s_d is the standard deviation for the fit itself. The results show values of A between about 14 and slightly less than 11, in rough agreement

TABLE 3. *Possible errors in least squares estimates*

Situation	Parameter estimates	Parameter variance estimates
A. Correct model, random errors in data		
Linear: correct weighting	Unbiased, efficient	Unbiased
Linear: wrong weighting	Unbiased, not efficient	Biased
"Linear": weighting of independent variable	Biased	Biased
Nonlinear	Biased	Biased
B. Correct model, systematic errors in data	Biased	Biased
C. Incorrect model	Systematic error	Systematic error

TABLE 4. *Least squares results in the absence of systematic error: exact 3DGE data fitted by the 3DGE model*

δ_i	Linear equation		Nonlinear equation	
	p-Weighting	V-Weighting	p-Weighting	V-Weighting
δ_0	($\hat{A}_0=-2.7\times10^{-13}$)	($\hat{A}_0=-2.4\times10^{-13}$)	-7.0×10^{-15}	-7.0×10^{-15}
δ_1	3.3×10^{-13}	3.0×10^{-13}	3.5×10^{-13}	2.9×10^{-13}
δ_2	-1.5×10^{-12}	-1.3×10^{-12}	-1.4×10^{-12}	-1.1×10^{-12}
δ_3	-1.9×10^{-11}	-1.7×10^{-11}	-1.9×10^{-11}	-1.5×10^{-11}
s_d	2.3×10^{-12}	2.0×10^{-14}	2.3×10^{-12}	2.0×10^{-14}

with the formula. There appears to be no significant tendency for the linear results to be better than the nonlinear ones, and one can scarcely conclude that much of the bias of table 3 is showing up here. In fact, bias is only important when random errors are present; in the A cases of table 3, bias approaches zero as the random error goes to zero. Incidentally, since δ_0 is zero in the linear case when V_0 is taken fixed at its exact value, A_0 is given in its place; since the true value of A_0 is zero, this is an absolute, not relative error.

The results of table 4 were calculated with $N=37$ points, covering the range $0 \leq z \leq 0.476$. Let the maximum value of z be denoted z_r. In earlier work [1], weighting of both the dependent and independent variable data values was discussed. The related standard deviations were denoted σ_V for the V variable and σ_p for the p variable. The p-weighting of table 4 takes $\sigma_p=1$ and $\sigma_V=0$ (weighting of dependent variable only), while the V-weighting uses $\sigma_p=0$, $\sigma_V=1$. In the linear equation case, the V-weighting chosen leads to somewhat different weighting of the actual independent variable x used [1]. Table 4 indicates slightly improved results for the V-weighting over the p-weighting, and no bias arising from V-weighting in the nonlinear situation is apparent. The differences between the s_d's for p and V weighting arise because the p-weighting s_d is a measure of the least-square fit residuals when they are all in pressure and is here in kilobars, while for V-weighting the residuals are all forced to be in reduced volume and s_d is then dimensionless. The ratio between the s_d's is roughly K_0.

Although we shall use the usual inaccurate Gaussian-elimination-with-pivoting solution of the least squares equations in the following, all inaccuracies introduced thereby are four or more orders of magnitude smaller than the systematic errors we consider. Such systematic errors in parameters and standard deviations are illustrated in table 5, where the 3DGE data are fitted with p-weighting using the incorrect BE_2 model. Results for δ_i and Δ_i are first given for two different ranges of z, from zero to ~ 0.048 and ~ 0.48. Note the strong increases in these error measures both with range and with the index i. Also included in the table are fitting results obtained for the complement range, all points contained in the second range but not in the first. As might be expected, the parameter estimates are somewhat worse for this coverage than for the largest span shown, even though s_d itself is somewhat better.

TABLE 5. *Least squares results showing systematic errors: exact 3DGE data fitted by the* BE_2 *model*

Range / i	$0 \leq z \leq 0.0476$ $N=19$		$0 \leq z \leq 0.476$ $N=37$		$0.0572 \leq z \leq 0.476$ $N=18$	
	δ_i	Δ_i	δ_i	Δ_i	δ_i	Δ_i
0	1.7×10^{-8}	2.4	1.1×10^{-5}	3.8	1.2×10^{-4}	9.7
1	-1.1×10^{-5}	5.8	-1.7×10^{-3}	11.0	-4.8×10^{-3}	14.0
2	4.5×10^{-4}	12.0	1.2×10^{-2}	21.0	1.9×10^{-2}	24.0
3	4.9×10^{-2}	29.0	2.7×10^{-1}	43.0	3.4×10^{-1}	45.0
s_d	1.31×10^{-6}		1.02×10^{-3}		4.62×10^{-4}	

All nonlinear least squares fitting in the present paper has been carried out using the Deming iterative method of solution [1, 11]. Although this is an approximate method [12], the resulting errors in the estimated parameter values are generally negligible. O'Neill et al. [12] have presented a more accurate iterative solution of the least squares problem with weighting of both dependent and independent variable, applicable only for polynomial equations. We have recently generalized and improved this solution so that it applies to equations of any form and converges much more rapidly [13]. This method, applied to the situations of table 5, yields essentially the same δ_i values as those in the table but Δ_i's some 25 to 40 percent larger than the tabular values. These increases thus mainly arise from smaller $(s_d)_\theta$'s produced by the new *least* squares solution. Although the new method leads to an essentially exact (in the sense of iterative convergence) least squares solution when round-off errors are negligible, the results cited above and those in the table show the presence of large systematic errors in δ_i and Δ_i arising from wrong model choice. The rest of the present paper is primarily concerned with δ_i's and p-weighting, where the differences between the Deming and improved estimates are negligible.

Although table 5 gives one some idea of systematic error effects, much more is provided by the least-squares results of figures 6 to 12. The same exact 3DGE data were fitted with the various other equations for four different ranges, all including $z=0$. The four values of z_r used were ~ 0.048, 0.143, 0.476, and 2.00, and the corresponding number of z values were, respectively, 19, 29, 37, and 58. All points used in a given fitting were included in the ones with larger z_r.

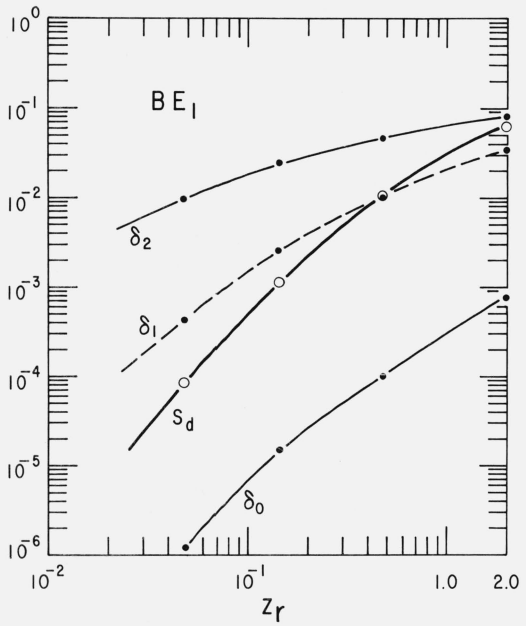

FIGURE 6. *Parameter relative errors,* δ_i, *and standard deviation of fit,* s_d, *versus fitting range* $(0 \leq z \leq z_r)$ *for least squares fitting of exact 3DGE data with the* BE_1.

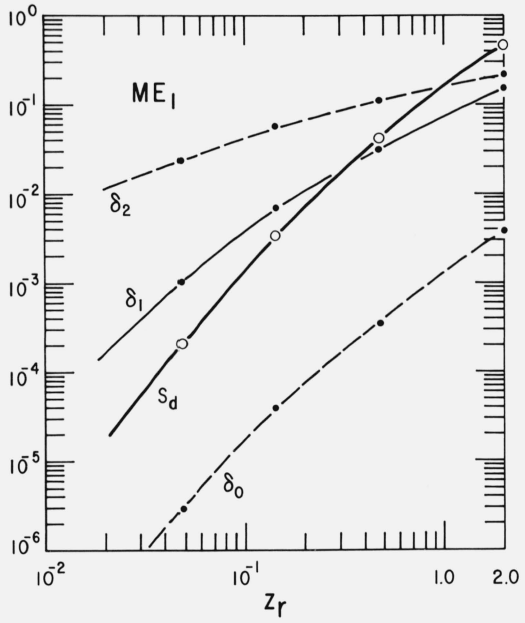

FIGURE 7. *Parameter relative errors, δ_i and standard deviation of fit, s_d, versus fitting range ($0 \leq z \leq z_r$) for least squares fitting of exact 3DGE data with the ME_1.*

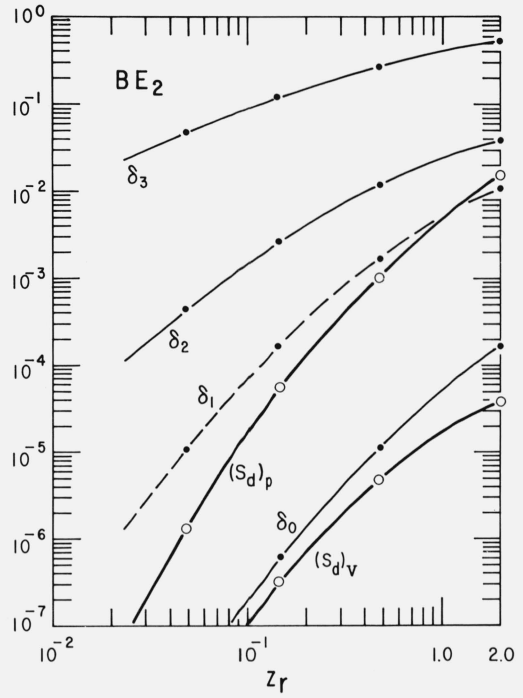

FIGURE 9. *Parameter relative errors, δ_i, and standard deviation of fit, s_d, versus fitting range ($0 \leq z \leq z_r$) for least squares fitting of exact 3DGE data with the BE_2.*

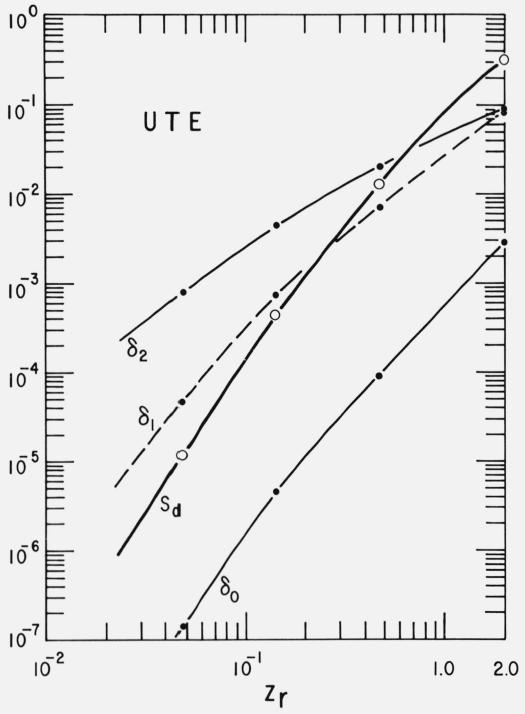

FIGURE 8. *Parameter relative errors, δ_i, and standard deviation of fit, s_d, versus fitting range ($0 \leq z \leq z_r$) for least squares fitting of exact 3DGE data with the UTE.*

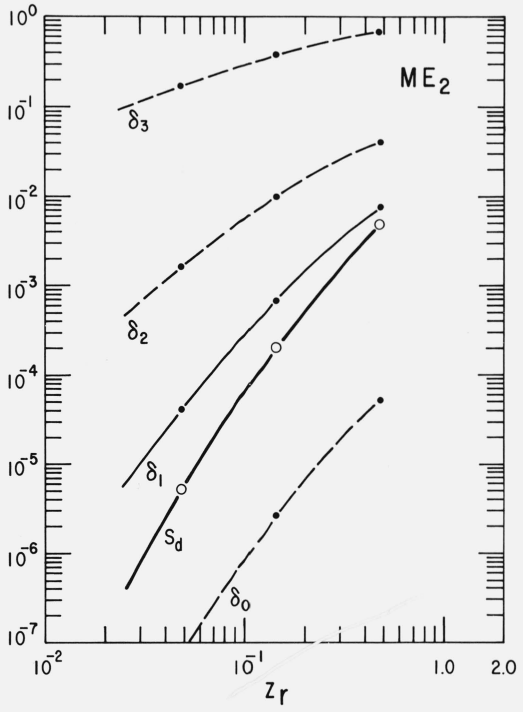

FIGURE 10. *Parameter relative errors, δ_i, and standard deviation of fit, s_d, versus fitting range ($0 \leq z \leq z_r$) for least squares fitting of exact 3DGE data with the ME_2.*

429-270 O - 71 - 4

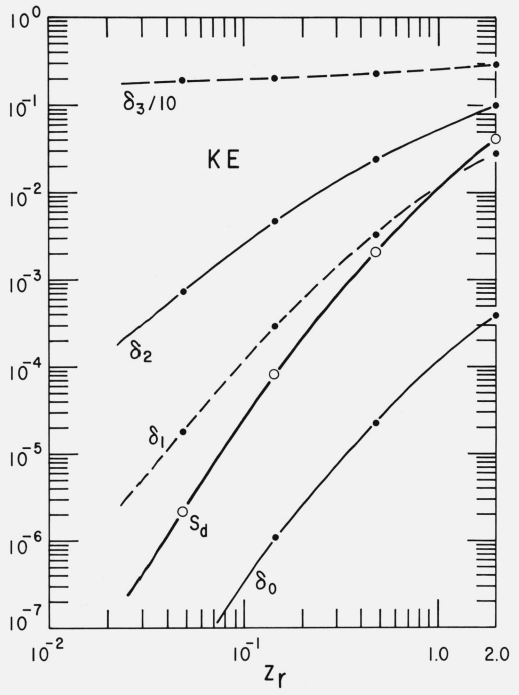

FIGURE 11. *Parameter relative errors, δ_i, and standard deviation of fit, s_d, versus fitting range ($0 \leq z \leq z_r$) for least squares fitting of exact 3DGE data with the KE.*

Figures 6 to 12 show how the systematic errors, represented by the δ_i's and by s_d, change as the fitting range is extended. As usual, dashed lines indicate negative values. All δ_i results shown were obtained with p-weighting: $\sigma_V = 0$, $\sigma_p = 1$. Results obtained with V-weighting were closely similar. Unlike the δ_i's, which are relative, the s_d's are absolute and, with p-weighting, measure the overall goodness of fit in pressure units, as already mentioned. Thus, for example, figure 6 indicates that s_d for the BE_1 fitting over the range $0 \leq z \leq 2$ is nearly 0.1 kbar. All s_d curves were obtained with p-weighting except the one marked $(s_d)_V$ on figure 9. Here we compare the s_d's obtained from p and V weighting. The $(s_d)_V$ values are somewhat more than K_0 times smaller than $(s_d)_p$ values here. Note that, as expected, $(s_d)_V$ and δ_0, the relative error in V_0, are quite close together over much of the range.

For V weighting, s_d is an overall measure of the residuals in V. Its absolute value in figure 9 at $z = 0.143$ of about 3×10^{-7} (the maximum magnitude of a volume residual was $\sim 6.5 \times 10^{-7}$) is about two orders of magnitude smaller than the $\Delta V \equiv V_{BE_2} - V_{DGE}$ of $\sim 5 \times 10^{-5}$ shown in figure 3 for the same z. But this last figure is that applying when the correct parameter values are used in both equations. As expected, the least squares parameter adjustment in the BE_2 fitting of the exact 3DGE data makes it difficult to conclude (without independent knowledge of parameter values) that the BE_2 is the wrong model, as it is here. With some random errors in the 3DGE model data, least squares fitting using the BE_2 and KE, for example, would again generally lead to results which wouldn't allow one to identify either the BE_2 or KE as an incorrect model, even though they both would be.

The results of figures 6 to 12 show that when the range is extended, relative errors in all the parameters increase when wrong models are used. Further, the higher-order parameters are more inaccurate than the lower-order ones for all the ranges shown. Not much added accuracy in the higher-order parameters can be obtained by reducing the range and, in practical cases where random error is present, generally no added accuracy will be achieved by such reduction.

Figure 10 stops with a z_r of 0.476 because the volume predicted by the ME_2 is negative for $z \geq 1.85$, precluding a meaningful fit with $z_r \simeq 2.00$. Note that δ_3 for the KE and 3SE is so large that its values must be divided by 10 and 100, respectively, to allow plotting on the present scale. For the 3SE, even δ_2 must be divided by 10 as well. These results illustrate an important general point. The figures show that the BE_2 and KE are the best least-squares simulators of the 3DGE model as far as s_d is concerned. Yet even for the relatively low z value of 0.143 ($p = 15$ kbar), $|\hat{\Psi}|$ is about 10 percent high for the BE_2 and $\hat{\Psi}$ is of even the wrong sign for the KE. The average residuals arising from systematic error would, when all in volume, be mostly less than 10^{-6} in magnitude here. Even for the best data currently available such small residuals would be obscured by random error. Thus we see that it is possible that two different equations, both wrong (as here) or

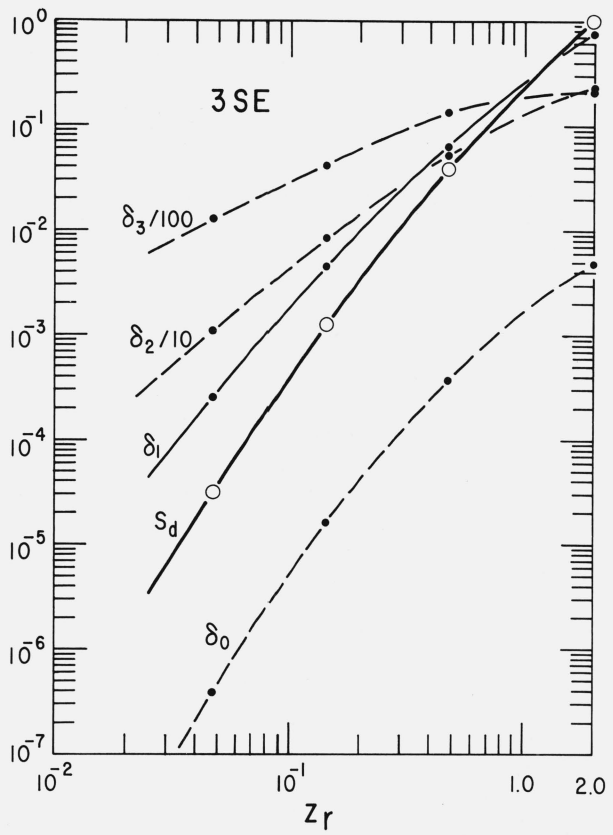

FIGURE 12. *Parameter relative errors, δ_i, and standard deviation of fit, s_d, versus fitting range ($0 \leq z \leq z_r$) for least squares fitting of exact 3DGE data with the 3SE.*

one wrong and one correct, may not be distinguishable by goodness of fit criteria, yet one may predict far better parameter values than the other. In the absence of other information, such as firm knowledge of the correct model or independent determinations of some of the parameter values, one will evidently always stand an appreciable chance of picking a "best" model which yields some quite poor parameter estimates. The better the accuracy of the data and the wider its range, the better the higher-order parameter estimates will be since the final model chosen will be forced to be closer to the correct model to achieve an adequate fit.

The monotonic increase of δ_i and s_d with fitting range illustrated in figures 6 to 12 is, of course, indicative of the use of an incorrect and inadequate model and is by no means limited to the equation of state area. In most if not all cases of practical interest, we may expect to find this sort of behavior: the wider the range used in least-squares fitting of a possible, "close," but still incorrect model, the greater will be s_d and the parameter error magnitudes. It should, however, be remarked that this conclusion only applies in the usual case where the model is not asymptotically correct as the range is extended indefinitely. The wider the range used, generally the more difficult it will be for an incorrect model to simulate the correct one.

This increase of errors with range may frequently be used in practical cases as a powerful means of discriminating against incorrect models. When random errors in the data are sufficiently small that the systematic errors arising from an incorrect model choice dominate s_d, it will generally increase with the fitting range, as illustrated here. Such an increase thus clearly signals an incorrect model choice for the range of data fitted. Since most models only apply adequately in any case over a limited range of a variable such as pressure or temperature, extension of the fitting range beyond the region of applicability of the best available model will always eventually result in an increase in s_d. Thus, in any least squares fitting where the range of applicability of the model used in unknown, extrapolation outside the fitted range of data should be approached with the utmost caution and avoided if possible.

The present paper deals with exact data and actual relative errors of parameters, but true parameter errors will not be available in a usual experimental situation. Nevertheless, when s_d increases because of the wrong model choice, the estimated parameter standard deviations will generally increase for the same reason. Thus, these quantities, ordinary results of a least squares fitting, may also be used along with s_d to help discriminate against an inadequate model.

There are some interesting general aspects to the present results obtained with least-squares fitting of the wrong model. The residuals (here given by observed values minus calculated values) show the following behavior: The number of runs (number of sign changes plus one), u, for the ME_1, BE_1, and UTE, for which $n=3$, is 4, while $u=5$ for the remaining equations for which $n=4$. The general result, $u=(n+1)$, is not very surprising but bears emphasizing. Further, the sign of the first residual run (which, together with knowledge of u determines the signs and order of all the runs) is specific to the equation considered. For the present fitting of 3DGE data, this sign is $+, -, -, +, -, -, +$ for the ME_1, BE_1, UTE, ME_2, KE, BE_2, and 3SE, respectively. The number of runs and their sign distributions were invariant in the present situation to the following: (a) p or V weighting, (b) the range of the data and its placement (all low p, all high p, all in the middle, etc.), and (c) the sign of Ψ. Even though not all extremes were investigated, this high degree of pattern invariance is likely to be quite general and may itself be of considerable usefulness in helping to distinguish models.

Although we have not done it, one could readily establish a matrix of first signs obtained using data calculated from one of the present eight specific equations and fitted with another one of the eight. Then, in practical situations where it was believed that the correct model was one of the eight, many possibilities could be quickly eliminated by comparison with the sign of the first residual run obtained on fitting the actual data. This would only work, of course, provided random errors were considerably smaller than systematic ones and hence didn't appreciably perturb the residual pattern. With many data points, considerable perturbation of this kind could be tolerated, however, since decisions could be made on the basis of a smoothed residual pattern rather than the actual noisy pattern.

A partial comparison of the above type has been made earlier for the ME_1 and UTE [14]. There, V_0 was taken fixed, so $n=2$. As expected, u was found to be three for both UTE fitting of exact ME_1 data and for ME_1 fitting of UTE data. The initial run signs were $+, -$, respectively, for the above two fittings.

5. Summary

This paper has been primarily concerned with discussing methods of discriminating between specific equations of state and has demonstrated considerable limitations on the possibility of adequate discrimination between "close" equations. We have found the somewhat surprising result that equations which cannot be adequately discriminated on the basis of least squares goodness of fit over even a wide pressure range may yet lead to estimates of higher-order parameter relative errors differing in sign and by an order of magnitude in absolute value for even a narrow pressure range, much less a wide one. The present methods, results, and conclusions can be generalized to a considerable degree to apply to model discrimination outside the equation of state area and are pertinent for linear models and for those nonlinear in their parameters, independent variable, or both. Thus, the following general conclusions, based on the present specific results, are likely to apply widely to the general data analysis field.

More than one mathematical model should usually be tested against the data in order to select, if possible, that model which fits best by objective criteria. As the range of data is progressively increased for which least squares fittings are carried out, the initial or eventual appearance of increases in s_d and $(s_d)_\theta$ indicates the

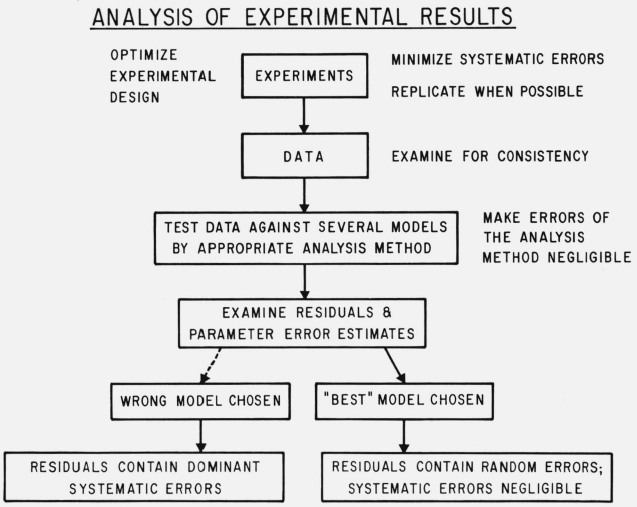

FIGURE 13. *General block diagram for data analysis.*

presence of systematic fitting error arising from an inadequate model choice. Such error will also usually show up in highly correlated residuals exhibiting, at least approximately, a number of sign-changes equal to the number of fitted parameters. The range of a causal experimental variable such as pressure, voltage, temperature, etc. should be increased to the maximum degree possible in order to allow the testing of a model for adequacy over the widest practical data range.

When two or more models have been found that represent the data over a given range with approximately the same goodness of fit and without signs of systematic errors from wrong model choice, it is still possible that one or more models may yield much better or much worse least-squares parameter estimates than the others. Additional independent information about likely parameter value ranges will usually then be necessary to allow a selection of the most appropriate equation for parameter estimation. Extrapolation of a given model-parameter value set beyond the range of data on which it is based is always dangerous.

When data smoothing or interpolation is the object, the possibility of discrimination between two models which yield equally good least squares fits to the data should be examined by the case B procedure of section 3. If the differences in dependent variables calculated with the same reasonable set of parameter values in each model are comparable to or smaller than the estimated random errors in the data, discrimination is impractical for that data set.

Figure 13 shows, in very diagrammatic form, appropriate steps in data analysis aimed at establishing a "best" model (including specific parameter value estimates). Some of the actually interrelated steps involved in this figure are presented differently in the flow chart of figure 14. This diagram is included for the benefit of those readers who may wish to apply the procedures discussed in this paper to other discrimination and parameter estimation problems.

For figure 14 we have assumed that a data set over a range, R_{max}, has been taken, and that we wish to test potential models over the maximum range if possible. The flow chart orders the tests as (1) case B, (2) runs, and (3) case A. If no models appear appropriate after the first series of tests, provision is made for decreasing the range of the data used in the tests in order to determine the acceptable maximum range for parameter estimation.

In the flow chart, we have abbreviated the test for case B by the notation $|\Delta y_{ij}| < \sigma_y$. Here we mean that all or nearly all of the deviations should be less than the expected errors in the data. Note that "nearly all" is appropriate because of the possible presence of random outliers. For the same reason, the test $u > n+2$ should also be considered approximate and applied judiciously. Note also that σ_y may vary with x (heteroscedastic case); the test should be so applied when appropriate. Other symbols introduced here are ϵ_s, defined to be the acceptable level for standard deviation of the least squares fitting, and ϵ_d, defined as the level below which standard deviations of two separate fits are indistinguishable.

Good data are usually expensive, yet too little adequate data analysis is the general rule. It is better to do too much such analysis than too little.

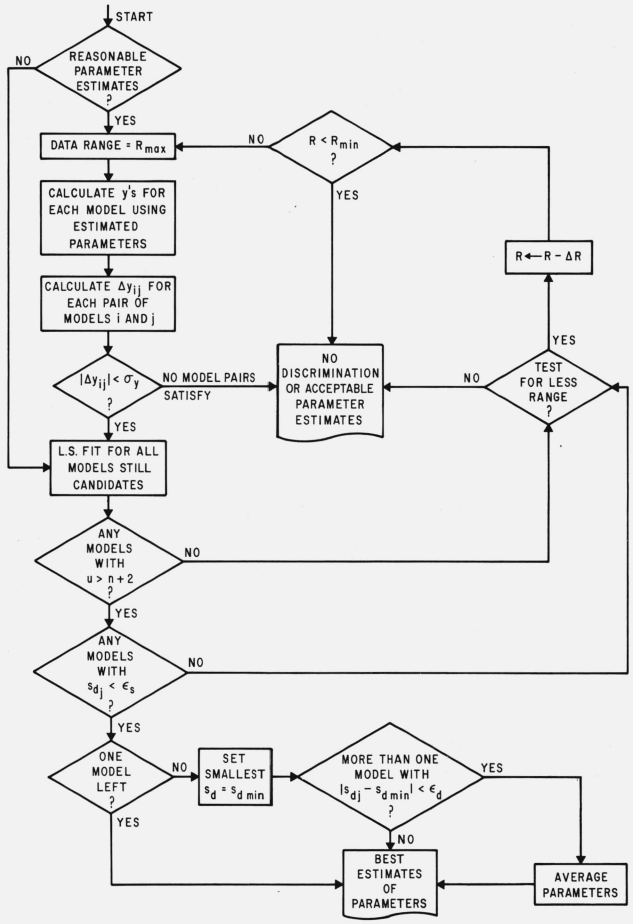

FIGURE 14. *Flow diagram for discrimination and parameter estimation tests.*

We much appreciate the very helpful comments of E. L. Jones and of a referee.

6. References

[1] Macdonald, J. R., Review of some experimental and analytical equations of state, Rev. Mod. Phys. **41**, 316 (1969).

[2] Davis, L. A., and Gordon, R. B., Compression of mercury at high pressure, J. Chem. Phys. **46**, 2650 (1967).

[3] Barsch, G. R., and Chang, Z. P., Ultrasonic and static equation of state for cesium halides, Proc. Symp. on Accurate Characterization of the High Pressure Environment, National Bureau of Standards, Gaithersburg, Md., Oct. 14–18, 1968. (NBS Spec. Publ. No. 326, pp. 173–186). The room-temperature CsI parameter values used in the present work were taken from a preprint of this paper distributed by the authors. Values appearing in the final version are $K_0 = 118.9 \pm 0.5$ kbar, $K_0' = 5.93 \pm 0.08$, and $K_0'' = -0.073 \pm 0.008$ kbar^{-1}. Note that this $|K_0''|$ is nearly three of its standard deviations larger than the value of -0.052 ± 0.002 kbar^{-1} presented earlier.

[4] Lazarus, D., The variation of the adiabatic elastic constants of KCl, NaCl, CuZn, Cu, and Al with pressure to 10,000 bars, Phys. Rev. **76**, 545 (1949).

[5] Ruoff, A. L., Linear shock-velocity-particle-velocity relationship, J. Appl. Phys. **38**, 4976 (1967).

[6] Monfort III, C. E., and Swenson, C. A., An experimental equation of state for potassium metal, J. Phys. Chem. Solids **26**, 291 (1965).

[7] Kell, G. S., and Whalley, E., The PVT properties of water. I. Liquid water in the temperature range 0 to 150 °C and at pressures up to 1 kbar, Phil. Trans. Roy. Soc. (London) **A258**, 565 (1965).

[8] Vedam, R., and Holton, G., Specific volumes of water at high pressures obtained from ultrasonic-propagation measurements, J. Acoust. Soc. Am. **43**, 108 (1968).

[9] Macdonald, J. R., Accelerated convergence, divergence, iteration, extrapolation, and curve fitting, J. Appl. Phys. **35**, 3034 (1964).

[10] Wampler, R. H., A report on the accuracy of some widely used least squares computer programs, J. Am. Stat. Assoc. **65**, 549 (1970).

[11] Deming, W. E., Statistical Adjustment of Data (John Wiley & Sons, Inc., New York, 1943).

[12] O'Neill, M., Sinclair, I. G., and Smith, F. J., Polynomial curve fitting when abscissas and ordinates are both subject to error, Computer J. **12**, 52 (1969).

[13] Powell, D. R., and Macdonald, J. R., A rapidly convergent iterative method for the solution of the generalized nonlinear least squares problem, submitted to Computer J.

[14] Macdonald, J. R., Some simple isothermal equations of state, Rev. Mod. Phys. **38**, 669 (1966).

(Paper 75A5–678)